Perceiving, Measuring and Modeling Materials

Christopher W. Tyler

Smith-Kettlewell Eye Research Institute

Abstract

This paper focuses in the interpretation of material properties of reflectivity and specularity assessed by the visual system under illumination consisting of both a focal and a diffuse component (the 'sun-and-sky' illumination assumption). This assumption provides for four kinds of luminance gradients: gradients of incident illumination, gradients of reflectivity, gradients of secondary self-illumination and gradients of shadowing. The analysis considers the dissociation of the material properties carried by specularity from the geometric properties of object shape, taking the sinusoidal surface as a canonical shape exemplar.

INTRODUCTION

Visual Properties of Materials

The perception of material through vision involves four primary forms of assessment: color, specularity, texture and transparency. If the material composing an object is perfectly smooth and opaque, its surface can only vary in gradients of luminance and chromaticity. These gradients, however can derive from four sources: gradients of incident illumination, gradients of reflectivity, gradients of secondary self-illumination and gradients of shadowing (self-shadowing from occluding parts of the same object and inter-shadowing from other objects in the case of complex scenes). The material property is specific to the reflectivity gradient – the other three components being properties of the illumination and the object geometry. Analytically, the net luminance L at any point is the product of the incident polar illumination function I(θ,ϕ) (taking into account the incident illumination function and the geometry of the object and surrounding objects) and the intrinsic reflectance function $R(\theta,\phi)$, such that

$$L = \int I(\theta,\phi) \cdot R(\theta,\phi)\, d\theta d\phi$$

In general, as emphasized by Barron & Malik (2013), inferring the reflectance function from the image luminance is a heavily underconstrained problem, particularly in view of the complexity of interreflections from the object geometry overlaid on the overall distribution of the incident illumination. Locally, however, the problem is somewhat alleviated by the fact that the local gradient is inherently one-dimensional in the direction of the 2D gradient maximum. For these reasons, the present treatment will focus on the one-dimensional problem of assessing the shape, illumination and reflectivity from the shading (or luminance) image, or what Barron & Malik (2013) dub the SIRFS problem. The visual system is this faced with an extraordinary degree of difficulty in performing this underconstrained reconstruction, which it does effortlessly for most visual scenes.

This feat is all the more remarkable because, as is well known, visual processing operates with a logarithmic transform of light intensity to perceive brightness, as depicted in Fig. 1. Typical computer monitors, whether CRT or LCD technology, are constructed with an accelerating gamma function with a power of 2 - 2.5, which would provide compression for a perceptual compression function with the inverse expansion. Such a gamma function would provide accurate correction for a perceptual compression function with the reciprocal power, but it should be noted that it is not fully accurate for correction for a *logarithmic* compression function. On the other hand, Stevens (1961) found that perceived brightness increased with a *power* function of 1/3 with luminous intensity, suggesting the perception was governed by a different form of law (Fechner's Law) than the logarithmic relation for increment thresholds (Weber's Law). Stevens' result would tend to validate the power function used for most monitors, although the actual power is substantially different from what Stevens reported, leaving a comparable degree of discrepancy in either case. Since there is no obvious

Proc. SPIE 9018, Measuring, Modeling, and Reproducing Material Appearance, 901802 (March 10, 2014); doi:10.1117/12.2046104

criterion for the veridical appearance, however, it is likely that neither discrepancy would be visually apparent (although they may leave an unquantifiable impression of artificiality in viewing the images on a monitor relative to the real objects).

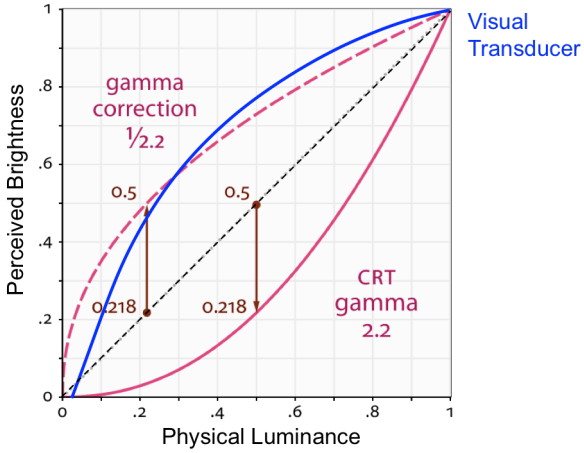

Fig. 1. Perceived brightness transducer (blue curve) compared with the power gamma correction (dashed red curve) compensated by the typical CRT gamma function (solid red curve).

One way of providing a visualization of the compression is to view luminance profiles with defined symmetry properties that allow direct assessment of the spatial configuration of the brightness gradients. For this purpose, sinusoidal luminance profiles have the appropriate property, since the dark and light halves of the cycle should have the same spatial curvature but of opposite sign. An example is reproduced in Fig. 2, based on the concept of a 5-cycle radial sinusoidal rosette. This rosette image form has three advantages:

1. It covers a wide range of spatial frequencies, allowing assessment of the perceived distortion over a wide range of spatial frequencies.
2. Its spatial symmetry enforces the assumption that the lighting is radially symmetrical, consisting of the combination of a radially symmetric diffuse component and a point source normal to the center of the image.
3. Its odd number of spokes is incompatible with alternating interpretations of the depth structure giving rise to the luminance profile.

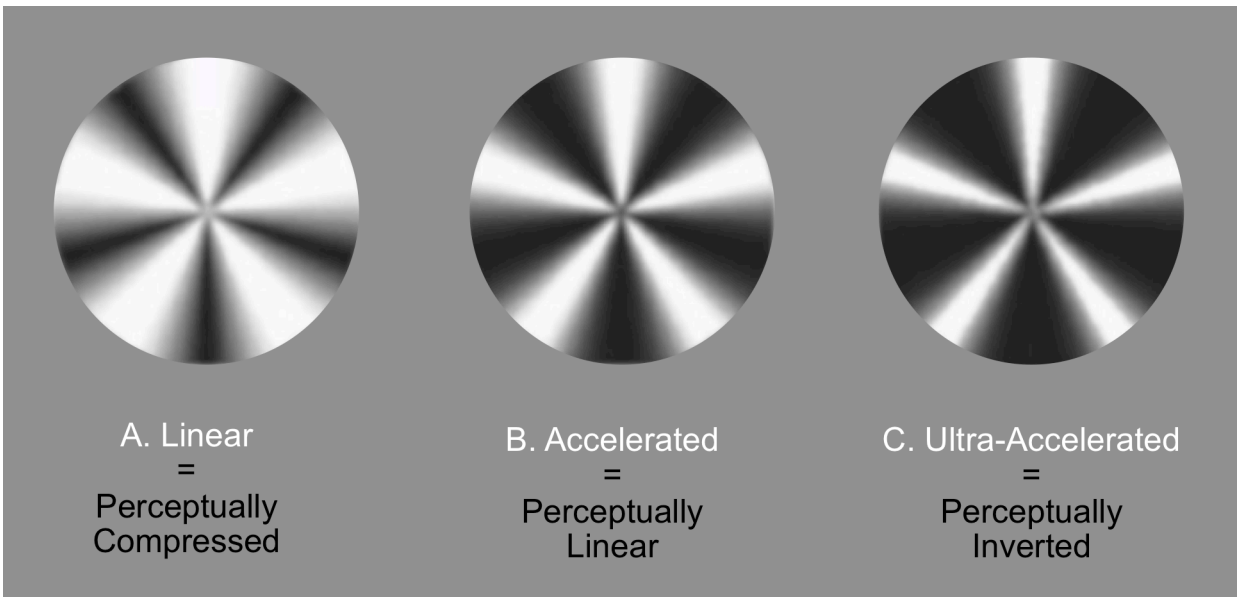

Fig. 2. Three examples of radial sinusoidal luminance images with different distortion functions: A: linear gamma that is perceived as compressive; B: accelerating gamma that is perceived as linear; C: ultra-accelerated gamma that is perceive as the inverse brightness function of A. These are approximate depictions of effects that were accurately developed on a monitor calibrated as linear to 99% accuracy.

Proc. SPIE 9018, Human Vision and Electronic Imaging XIX, 90180A (March 10, 2014); doi:10.1117/12.2046104

Fig. 2A demonstrates the profound brightness compression experienced with a sinusoidal profile that is fully linear in terms of its luminous intensity. The bright regions appear broad and smooth, while the dark regions appear to drop down to a sharp line. The mid-gray dividers appear to occur at only about a quarter of the cycle width, supporting a compelling degree of brightness compression. Note that the form of the brightness profile looks similar (i.e., the radial gradients appear straight) most of the way into the center of the rosette, implying that the compression is invariant with spatial frequency. In the center the present images are grayed out by a slight blur, but on the linearized monitor they remained constant all the way to the center, implying that the compression is a point-wise nonlinearity, not some function of lateral inhibition operating at low spatial frequencies.

Incidentally, it is notably that the perceived contrast does not decrease at low spatial frequencies in the radial displays, as may be expected from the threshold sensitivity functions. This is a property that has been described before for high-contrast images by Georgeson & Sullivan (1975), who explained it in terms of renormalization narrowband spatial frequency channels according to Bayesian constraints derived from past experience. This explanation, however, seems implausible in light of the evidence that there are no low-frequency channels (Blakemore & Campbell, 1969; Kontsevich & Tyler, 2012), the range probed by the rosettes of Fig. 2. Although the lowest-frequency channel varies with eccentricity from the fovea (Kontsevich & Tyler, 2012), it is evident that the appearance of the rosette does not change with fixation at different locations, presenting different spatial frequency content to the fovea (where the range not covered by spatial-frequency channels is the most extensive, and the inhibitory contrast reduction would be expected to be the greatest). The implication is that the low-frequency fall-off from the lowest-frequency channel most be compensated by some mechanism that operates as a spatial integrator.

SIRFS Perception

The SIRFS problem of the reconstruction of the surface structure (S), illumination properties (I) and material reflectance function (R) from the shading gradient function is an issue for the perceptual system just as it is for computer vision. Fig. 2 illustrates that the visual system can derive stable estimates of all three functions from a single one-dimensional (radial) function. The brightness appearance has been discussed. What surface structure and material properties does it appear to have? Viewed as an object, it is seen as a physical ('horse-show') rosette, nearer in the bright regions and further in the dark regions. This may seem obvious, but note that this is not a property that could be conveyed by a (centered) point source illumination. Such illumination would produce frequency doubling of the luminance profile relative to the surface shape (Tyler, 1998), where the near and far regions would both be bright and the intervening gradient regions would be darker. The appearance of the bright regions near and the dark ones far is consistent only with the assumption of diffuse illumination uniformly from all directions, or at least symmetrically about the center of the rosette.

This problem is still triply underconstrained, and to make headway on it, the issue will be further simplified according to a 'sun-and-sky' assumption by subdividing the illumination function into a point-source and uniform-diffuse component. Most computer graphics scene rendering and reconstruction techniques assume point source illumination from some asymmetric angle. Langer & Zucker (1994) and Tyler (1998) took the alternative approach of assuming uniform diffuse illumination, the 'sky' component that is almost always present due to secondary reflections either in the atmosphere or from the surfaces of enclosing structures. One can make the case that the diffuse component is the more important of the two, and deserves to be considered as a neglected default assumption of the typical source of illumination in natural images of the world we inhabit.

The third aspect, the material property of the rosette, is also evident in Fig. 2A. Seen as a surface, it appears as a matte, chalky material. That is, the material property is assumed to be Lambertian. The other two rosettes in Fig, 2 B,C are provided to illustrate the consequences of variations in the luminance profile. When the gamma is accelerated to provide an approximately sinusoidal brightness profile (Fig. 2B), the surface is seen as sinusoidal in shape. However, the material is now perceived as partially shiny, like brushed aluminum. With further acceleration of the gamma, the brightness profile is approximately inverted from the linear appearance, with narrow bright lines and broad dark regions. The material, however, is now perceived as shiny while the shape remains approximately sinusoidal. It becomes apparent that the narrow bright regions conform to the option of partially specular highlights, allowing the surface to relax to a less sharp form, an

Proc. SPIE 9018, Measuring, Modeling, and Reproducing Material Appearance, 901802 (March 10, 2014); doi:10.1117/12.2046104

option that is not available for the shape determination from the narrow dark regions in fig. 2A. It is this scission between the shape interpretation and the highlight assessment from a single one-dimensional luminance profile that allows the visual system to assess the material property and the shape property simultaneously, apparently on the basis of a Bayesian smoothness constraint on the shape interpretation.

Diffuse and Focal Illumination as the Default Assumption for SIRFS Reconstruction

It is evident, however, that the input to the triple SIRFS interpretation is the luminance profile and that the problem faced by the visual system is to reverse-engineer the sequence of surface shape/incident illumination/reflectance function/brightness compression that gives rise to the luminance information impinging on the retina, as diagrammed in Fig. 3. In this regard, one can ask why the visual system would have evolved to provide such a strong brightness/luminance compression function, when it obviously results in such a pronouncedly non-veridical percept of the luminance function. The answer that will be developed in the following analysis is that the geometry of diffuse illumination is such as to form an expansive nonlinearity for the relationship between the surface shape and the luminance profile produced by diffuse illumination. Incorporating a simple brightness/luminance compression has the effect of providing a one-step shortcut for the full recovery of the forward sequence of Fig. 3. By interpreting the compressed brightness profile as a direct readout of the shape profile, the visual system can obtain a good approximation to the surface shape, under the twin assumptions of diffuse illumination and Lambertian shading (Tyler, 1998).

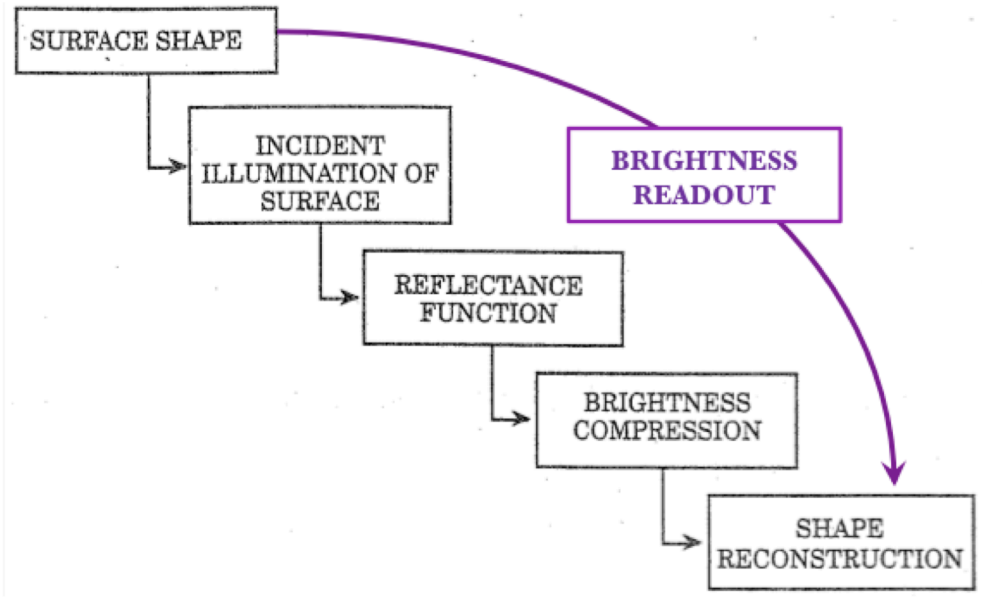

Fig. 3. Sequence of processing steps that need to be reconstructed in order to reconstruct the surface shapes in the world from the brightness image perceived in the brain, and simplifying shortcut (curved arrow).

For the present treatment, this concept can be expanded by one step to cover the analysis of the material property. The reflectance function is a property of the surface material, and hence the output of the brightness compression stage is treated as a signal to be split into a diffuse Lambertian component D = log(S) and a focal specular component $F = S^p$ (corresponding to the effects of the diffuse and focal aspects of the incident illumination). We may assume that the two components are independent processes and hence sum additively. Thus, the reconstruction problem faced by a visual system, human or computational, is to perform the scission operation of assigning the value of the proportionality constant k between the two components in the equation:

$$\log(L) = k \cdot S + (1 - k) \cdot \log(F)$$

Proc. SPIE 9018, Measuring, Modeling, and Reproducing Material Appearance, 901802 (March 10, 2014); doi:10.1117/12.2046104

and hence that the shape/highlight scission estimated by the human visual system, which encodes brightness as the logarithm of luminance, is described by

$$B = k \cdot S + (1 - k) \cdot p \cdot \log(S)$$

Surface Illumination Geometry

To illustrate the principle of surface reconstruction from the resulting brightness profile such as those of Fig. 2, consider the case of a one-dimensional sinusoidal surface. The lefthand panel of Fig. 4 shows three forms of illumination of a sinusoidal surface (A), the luminance profile for point source illumination at a grazing angle to the maximum surface angle (B), point source illumination normal to the mean surface (C), and fully diffuse illumination (D). The grazing angle direction (A) is the only case when point source illumination approximates a sinusoidal reflectance profile at the same cycle length as that of the surface, though it is important to note that it is phase shifted by 90° from the surface waveform, and is also shows a compressive distortion from a sinusoidal form. At steeper angles of incidence, all point source reflectance profiles have a frequency-doubled characteristic – perfectly sinusoidal when the angle of incidence is actually normal (C). The luminance peaks at the points of zero slope and is minimal at the points of maximum slope of the surface sinusoid. Thus, to reconstruct the surface sinusoid, the image processing would have to assign the bright regions alternately to the near and far peaks, which would be literally impossible for the five-spoked rosettes of Fig. 2 because such and alternating assignment requires an even number of peaks. The diffuse illumination from all directions (D) illuminates the surface in proportion to the viewing aperture of the 'sky' from each point on the surface together with second-order self-reflection terms (Tyler, 1998), with the net result of producing a waveform that is an expansive distortion of the sinusoidal surface waveform.

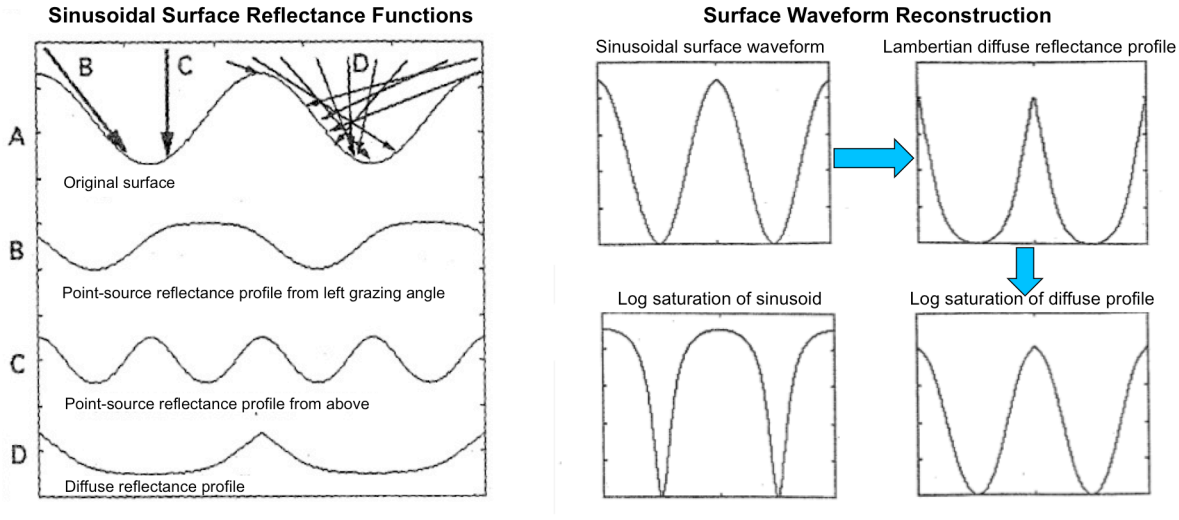

Fig. 4. Left: A. three forms of illumination of a sinusoidal surface; B. the luminance profile for point source illumination at a grazing angle to the maximum surface angle; C. point source illumination normal to the mean surface; D. fully diffuse illumination. Right: inference of sinusoidal Lambertian surface shape from log saturation of the luminance profile under the assumption of uniform diffuse illumination (D).

The right panel illustrates the reconstruction of the inference of sinusoidal Lambertian surface shape from logarithmic saturation of the luminance profile under the assumption of uniform diffuse illumination (D). The diffuse illumination produces a luminance function with narrow bright peaks (upper right panel). The two lower panels at right show the log brightness compression of the sinusoidal surface waveform and the diffuse reflectance profile, respectively. It is evident that the log compression provides approximate compensation for the distortion of the diffuse reflectance profile to regenerate the sinusoidal from of the surface waveform. Thus, under the particular assumptions of Lambertian reflectance and diffuse illumination, the logarithmic compression has the required effect of providing an effective shortcut for reconstructing the surface profile from the brightness function.

Proc. SPIE 9018, Measuring, Modeling, and Reproducing Material Appearance, 901802 (March 10, 2014); doi:10.1117/12.2046104

Conclusion

This treatment has focused on the perception of material properties through the luminance component of color and specularity. A host of other material properties are carried by aspects of texture, and also by transparency. Both are large topics in their own right, but the visual analysis of transparency may be viewed as a dissociation between the specular and reflectance aspects of an object image. If the optics of the transparent object are clear and undistorted, any transparency is undetectable without a specular component to the reflectivity. However, the specular component derives from different and hence completely dissociated sources of object shape and external reflections than those of the transparently visible scene. The specularity of a non-transparent object, on the other hand, depends on its material property of surface smoothness and is fully integrated with its shape.

References

Angelopoulou E, Lee SW, Bajcsy R (1999) Spectral gradient: a material descriptor invariant to geometry and incident illumination. Proc Seventh IEEE Int Conf Computer Vision, 861-867.

Barron JT, Malik J (2013) Shape, albedo, and illumination from a single image of an unknown object. 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 334-341.

Blakemore C, Campbell FW (1969) On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. J Physiol. 203:237-60.

Georgeson MA, Sullivan GD (1975) Contrast constancy: deblurring in human vision by spatial frequency channels. J Physiol. 252(3):627-56.

Kontsevich LL, Tyler CW (2012) A simpler structure for local spatial channels revealed by sustained perifoveal stimuli. J Vis. 13(1):22. doi: 10.1167/13.1.22.

Stevens SS (1961) To honor Fechner and repeal his law: A power function, not a log function, describes the operating characteristic of a sensory system. Science. 133:80-6.

.